REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	27 Dec 95	

4. TITLE AND SUBTITLE

Team Decision Making and the Tower of Hanoi: The Effects of Gender and Practice
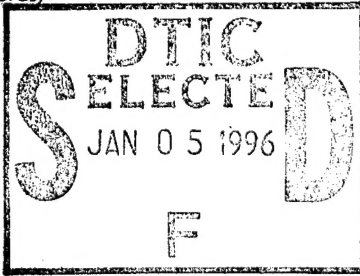
5. FUNDING NUMBERS

6. AUTHOR(S)

John M. Barker, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

AFIT Students Attending:

Florida State University

8. PERFORMING ORGANIZATION REPORT NUMBER

95-028D

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

DEPARTMENT OF THE AIR FORCE
AFIT/CI
2950 P STREET, BLDG 125
WRIGHT-PATTERSON AFB   OH   45433-7765

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for Public Release IAW AFR 190-1
Distribution Unlimited
BRIAN D. Gauthier, MSgt, USAF
Chief Administration

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

DTIC
SELECTED
JAN 0 5 1996
F

19960104 135

14. SUBJECT TERMS

15. NUMBER OF PAGES

73

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

DTIC QUALITY INSPECTED 1

THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

TEAM DECISION MAKING AND THE TOWER OF HANOI:

THE EFFECTS OF GENDER AND PRACTICE

By

JOHN M. BARKER, JR.

A Dissertation submitted to the
Department of Psychology
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:

Spring Semester, 1996

The members of the Committee approve the dissertation of John M. Barker, Jr. defended on December 1, 1995.

George E. Weaver
Professor Directing Dissertation

Marcy P. Driscoll
Outside Committee Member

John C. Brigham
Committee Member

K. Anders Ericsson
Committee Member
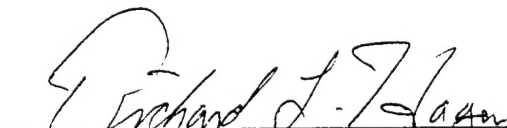
Richard L. Hagen
Committee Member

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Two experiments were conducted investigating the effects of gender and practice on team decision making while solving the Tower of Hanoi (TOH). All participants were undergraduates. The first experiment demonstrated that relevant practice on a 4-disk TOH wooden model, as compared to irrelevant practice using an Eight Puzzle, improves performance of individuals (20 females and 20 males) on a computerized 5-disk TOH and that males were faster than females at solving the TOH. The second experiment involved 84 two-person teams (84 females and 84 males) who solved the TOH in a design combining relevant practice gender (female, male, or neither) with gender mix (female-female, male-male, or female-male) resulting in 7 different treatment conditions (12 teams in each group). All team communications were audio-recorded and move information was recorded on the computer. The results suggest that teams with males receiving relevant practice solve the puzzle in less overall moves and in a shorter overall time as compared to teams with females receiving relevant practice. Results from the analysis of the search space data suggest that the teams with females who have received relevant practice reverse more moves compared to teams with males receiving

relevant practice. The verbal data indicate that males make more commands than females regardless of who receives relevant practice.

These data were interpreted in the context of several theoretical approaches to the study of team decision making. Only 8 out of a possible 48 participants revealed to their teammate that they had some expertise (prior practice) in solving the TOH puzzle. This result is consistent with some of the aviation accident reports that indicate a crewmember had critical information or expertise that was not shared with the team and therefore a poor team decision was made. The application of the current methodology to more naturalistic situations is discussed as well as the limit of the current findings to real-world settings.

TEAM DECISION MAKING AND THE TOWER OF HANOI:

THE EFFECTS OF GENDER AND PRACTICE

Decision making research traditionally emphasizes how individuals make decisions and what factors impact the quality of their decisions. Little research has been conducted on team decision making. Do teams use the same processes as individuals when making decisions? What factors impact team decision making?   When teams consist of individuals from various cultural backgrounds or various combinations of gender (all males, all females, or some of both) will the decision processes and outcomes differ?   The purpose of this research was to develop a methodology to study team decision making in realistic situations while at the same time answering some of the questions concerning factors affecting team decisions.

Two recent events have highlighted the importance of team decision making research. A Korean airliner crashed while attempting to land on the island of Cheju near the Republic of Korea ("Aircraft Crash," 1994). The initial details of the accident indicate that the pilot, a Canadian, and the copilot, a Korean, were arguing over the decision

1

to land. The pilot was making the landing and the copilot thought the aircraft was going too fast for a safe landing and demanded that they go around for another landing. The pilot did not agree and continued the landing as the copilot pulled on the controls to abort the landing -- resulting in a crash at the end of the runway. Fortunately everyone was evacuated before the aircraft was engulfed in flames. This is an example of poor team decision making. Is it possible that the "cultural mix" of the crew was a contributing factor in the accident? The individuals in another example of poor decision making were not so lucky. A federal report implicated poor judgment by crews and supervisors in the deaths of 14 firefighters ("Report Blames," 1994) battling blazes in the mountains of Colorado. Apparently they did not exercise sufficient caution and placed themselves in a deadly situation.

On July 19, 1989 a McDonnell Douglas DC-10 made an unscheduled crash landing at the municipal airport in Sioux City, Iowa. The captain, Alfred Haynes was on a routine flight for United Airlines when the aircraft suffered a catastrophic engine failure at an altitude of 37,000 feet and lost all flight controls. The actions by the captain and crew enabled an aircraft with 295 people on board that was virtually uncontrollable by any normal means, to make it

to an airport and save 184 lives. Captain Haynes attributes

the outcome to luck, communications, preparation, execution,

and cooperation (Haynes, 1991). I would add that

outstanding individual and team decision making was also a

factor in this crew's success. In contrast to the previous

examples this is an illustration of effective team decision

making.

A definition of team decision making and how it differs

from some of the group decision making research is necessary

before discussing some of the emerging theories in this

area. Orasanu & Salas (1992) define team decision making...

as the process of reaching a decision undertaken by

interdependent individuals to achieve a common goal.

What distinguishes team decision making from individual

decision making is the existence of more than one

information source and task perspective that must be

combined to reach a decision. While ostensibly working

toward the same goal, participants may have differing

agendas, motives, perceptions, and opinions that must be

melded into the shared product. (p. 328)

We must also consider the differences between teams and

groups. There has been an abundance of research on group

decision making, some of which may apply to team decision making. "Teams consist of highly differentiated and interdependent members; groups, on the other hand, consist of homogeneous and interchangeable members, like juries (Orasanu & Salas, 1992, p. 328)." Starting with these definitions of teams and team decision making, it is possible to review some of the research and its implications.

Some of the most promising research in this area involves the notion of shared mental models (Orasanu, 1990). An example of this concept involves crewmembers flying an aircraft, articulating situation assessment and possible solutions to an emergency, thus building a shared (between crew members and/or ground personnel) mental model of the situation. "Shared models assure that all participants are solving the same problem and create a context in which all can contribute efficiently" (Orasanu, 1990, p. 15). Specifically the crew/team should share an understanding of the definition of the problem, possible plans and strategies to solve it, interpretation of the cues and information, and roles and responsibilities of all involved. When team members can correctly predict other members' behavior and can anticipate their teammates' information needs (even

without verbal communication), then an accurate team mental model exists (Cannon-Bowers, Salas, & Converse, 1990).

Team process behaviors refer to specific behavior characteristics of successful teamwork. These behaviors include; effective communication, coordination, compensatory behavior, mutual performance monitoring, exchange of feedback, and adaptation to varying situational demands. Once researchers identify the requisite knowledge and skills for each of the team process behaviors, this information should allow us to better understand the phenomenon of team decision making and how to improve team decision making. This model requires specialized communication between members.

Orasanu (1990) examined the verbal protocols from two-person crews flying a Boeing 737 flight simulator at the NASA-Aimes research center. The flight scenario involved some in-flight emergencies coupled with bad weather and thus necessitated excellent crew decision making to complete the mission. The crews were divided into high and low performing crews depending on the number of errors (determined by domain area experts) committed during the flight. The crew utterances were categorized as situation awareness, information requests, ground communication, and the number of alternatives considered. There were

significant differences between the high and low performing crews concerning the amount of information requested and in their communication with ground controllers. The high performers requested three times more information than the low performers and talked to ground controllers 50% more. The data from this study suggest that the more complete the shared mental model, the better the crew decisions are.

The original study which provided the data for Orasanu's (1990) communication analysis was designed to study the effects of fatigue on aircrew performance (Foushee, Lauber, Baetge, and Acomb, 1986). The counterintuitive results showed that fatigued crews outperformed non-fatigued crews. This outcome becomes more understandable given that the fatigued crews had just recently flown an actual mission together, where the other crews had not. In other words the fatigued crews were more familiar with each other than the non-fatigued crews. It was suggested that flying together allowed crews to develop interaction patterns that facilitated building of a shared mental model, therefore improving their ability to make decisions and actually overcoming what might be expected to be a negative factor -- fatigue.

More recent research reviewing military aircraft accidents (Woody, Mckinney, Barker, & Clothier, 1994) and a

similar flight simulation study (Barker, Clothier, Woody, Mckinney, & Brown, in press) support Foushee's findings but only when crews have been together for a limited amount of time. There is evidence that when crews are together for more than 6 months as a team, maladaptive shared mental models or what is referred to in the group decision making literature as groupthink (Janis & Mann, 1977) may develop which paradoxically could decrease a crew's ability to make effective decisions.

While many team decision making phenomena can be explained by the concept of shared mental models, it is limited to certain team functions. Thordsen and Kleins' (1989) concept of a team mind is a more encompassing theory that parallels concepts from cognitive psychology explaining individual cognition. The model postulates that the mind of a team is analogous to the mind of an individual with three levels of the team mind: the behavioral level (overt actions), collective consciousness (reflected in communication), and subconsciousness (individual knowledge or interpretations that are not shared with others in the group).

According to a team mind framework, any action comes from the team which is considered a single entity. So, when an aircrew receives a communication from the ground

controllers that they are cleared for landing, it is the crew that lands the aircraft, not just the captain. Data relating to the collective consciousness of a team are collected from verbalizations of the team, usually by videotaping. The team's collective consciousness is therefore directly accessible while the team is performing, and studying this level will not interfere with the task. Additionally, knowledge held by one individual but not shared with the team can also be obtained from videotapes or interviews and provides insight into why "crew" errors are committed. The collective consciousness is similar to Orasanu's shared mental models.

The common concepts from cognitive psychology of limited attentional resources, working memory, workload, automaticity, and metacognition are used by Thordsen and Klein (1989) to explain and understand how teams make decisions. Since the team mind has a limited attentional capacity not everything should enter the team's collective consciousness. Many tasks could and should take place at the preconscious level. Metacognition is an important part of the team mind. Teams can record critical information to help avoid memory overload and teams can use appropriate strategies to avoid loss of information during interruptions.

Team mind errors occur which are different from the
skill/knowledge errors of individuals. Team errors result
from dysfunctional interactions among the three components
of the team mind. The team mind has intent which usually
comes from the leader and does not always need to be
explicit. Team members assume they understand the team
intent until cues indicate otherwise. The concept of
automaticity helps to explain how team decision making
becomes more effective and efficient as more of the tasks
are performed at the preconscious level.

The team mind model is based on several studies of team
decision making in operational environments (Thordsen &
Klein, 1989). The situations all involved teams making
decisions under high uncertainty of critical information,
with high risk consequences and under time pressure. The
first situation mentioned previously was a commercial
airline full-flight simulation with multiple malfunctions
(small-team structure). In addition a large-team structure
was analyzed in the context of information from quick-
decision exercises at the Advanced Warfighting class at the
Army Command and General Staff College. Teams of 15
individuals (rank of Major) operated as the general staff of
an army division. Another large-team consisted of students
from the National Defense University who were making crisis

decisions (an exercise) in response to geo-political events
which would be briefed to the NSA and then submitted to the
President. A final research project was conducted at the
Department of Energy's Central Training Academy involving
medium-large teams (12 individuals) that operated as a
Crisis Management Team in response to a hypothetical
terrorist attack or natural disaster. These research
projects involving command-and-control team performance
formed the basis for the team mind model and have
illustrated the usefulness of this model in explaining team
performance in teams that vary from a few members to a dozen
or more.

There are many other factors besides team size that
impact team decision making. The team mind analysis would
suggest that an important element in successful decision
making is the extent to which a team member is able to bring
relevant strategies to the decision process. One purpose of
the current study was to determine how information acquired
by one team member through prior practice is incorporated
into the decision process. Thus, individual team members
received relevant or irrelevant practice prior to a
structured problem-solving setting. A traditional problem-
solving task was selected since many of the previous
examples of decision making are also examples of problem

solving. For example, most of the critical decision making situations in airline cockpits occur as a result of a malfunction (or problem) with the aircraft systems or operating environment. The crew decides as a team how to best deal with the problem. The final outcome or goal may be well-defined such as shut-down an engine and land or it may be ill-defined such as save as many lives as possible. Regardless of the goal, team decisions must be made and there are many factors (mentioned earlier) which may impact the quality of these decisions. Although the use of such a standardized problem solving task may lack broad ecological significance, it does make it possible to compare results with those obtained in prior research.

The Tower of Hanoi puzzle has been used in a variety of ways to understand the processes involved in problem solving and decision making. This puzzle involves three vertical pegs or posts and a number of disks that vary in size which fit on the pegs. Initially, all of the disks are arranged on one of the pegs in a pyramid fashion (largest on the bottom). The task is to move all of the disks to one of the other pegs. The rules are: (a) only one disk may be moved at a time, and (b) a disk may not be placed on top of a smaller disk. Verbalization during practice trials and during the criterion trials has been shown to increase

performance on the task (Ahlum-Heath & Di Vesta, 1986).
Several studies have also shown a difference in performance
between males and females (Leon-Carrion et al., 1991). The
majority of research utilizing the Tower of Hanoi puzzle has
focused on individual problem solving and decision making,
not team decision making. The primary goal of the current
study was to investigate team decision making while solving
the Tower of Hanoi puzzle.

A second purpose of this study was to study how the
gender of team members influences information exchange and
problem-solving performance. There have been many studies
attempting to understand the differences between males and
females in group interactions, but none of these studies has
specifically investigated team decision making. Twenty
years ago this was not an issue in areas such as airline
cockpits because there were no mixed gender crews or all
female crews. Currently mixed gender teams can be found
making critical decisions and solving difficult problems in
many real-world settings ranging from airline cockpits to
hospital operating rooms. Past research has shown that
males and females differ in their social interactions (Shaw,
1981) which may impact group behavior and team performance.
Shaw also suggests that as differences between expected sex
roles in our society narrows, the differences in social

12

interactions may disappear. Many of the findings on mixed and non-mixed teams are based upon research which was conducted during the 1960s and 1970s (Morgan & Lassiter, 1992). "Because of the ever-changing sex roles in society, new studies are needed in order to evaluate the effects of gender mixes on team performance within the context of today's social milieu" (Morgan & Lassiter, 1992, p. 79). The current study examines the effect of gender on team performance during a problem solving task thus updating the literature with respect to presumed changes in gender roles in our society.

Previous research on the Tower of Hanoi indicates that females take longer than males to solve the puzzle. Based on this literature the results of the present study were expected to show whether this is true also for teams composed of all females as compared to male teams. The social interaction literature also supports the hypothesis of an interaction between gender mix and which gender receives relevant practice. Specifically, in the mixed gender conditions, the groups in which the male receives the relevant practice could be expected to out perform the groups in which the female receives relevant practice. Another question addressed involves the possible different strategies used by the teams of different gender make up.

13

This study also attempted to apply verbal protocol analysis to the study of team decisions made during problem solving. In addition, the model of the team mind was used as a framework to analyze team performance. Specifically the three aspects of the team mind -- behavior, consciousness, and subconsciousness were studied to determine the effect of team gender mix and variations in prerequisite knowledge and skills on the decisional processes and outcomes. In addition to an analysis of the performance data, protocols of two-person teams solving the Tower of Hanoi were analyzed using the methodology developed by Ericsson & Simon (1993) as well as traditional communication research methodology.

Two experiments were conducted examining the effects of gender receiving relevant practice and team gender mix on solving a 5-disk computerized version of the Tower of Hanoi. The first experiment used individuals as opposed to teams to validate that practice on a wooden 4-disk model of the Tower of Hanoi would improve performance on the computer version and to determine any gender effects on individual performance on the puzzle. Based on the results of the first experiment, the second experiment was conducted to investigate factors influencing team performance in solving the Tower of Hanoi.

Experiment 1: Validation Study

Purpose

The purpose of the first experiment was to provide direct evidence that a practice task would result in improvement on final task performance and to determine if there are any gender effects between individuals solving the Tower of Hanoi puzzle. The experimental task was a computerized version of the Tower of Hanoi puzzle (5-disk) and participants received prior practice on either a 4-disk wooden model of the Tower of Hanoi (relevant practice) or the Eight Puzzle (irrelevant practice). A between subjects design with 10 participants in each group was used.

Method

Participants. Forty upper division university students (20 women and 20 men, mean age = 21 years) volunteered to participate. Volunteers received "extra credit" in a cognitive psychology course for participating. All participants read and agreed to the conditions in the attached (see Appendix A) informed consent approved by the Florida State University human subject committee.

Apparatus. A wooden model of the Tower of Hanoi consisting of three dowels connected to a rectangular base was used for all relevant practice conditions. This model was a 4-disk version of the Tower of Hanoi. Two different

15

Eight Puzzles were used in the irrelevant practice
condition. One Eight Puzzle formed a picture of a dinosaur
and the other a cat (Appendix B). The Eight Puzzles consist
of eight blocks which can slide into an open space in a 3 x
3 matrix. The pattern is scrambled and the goal is to
rearrange the blocks to recreate the picture.

A computer version of the Tower of Hanoi called the
stack of blocks was implemented using Toolbook software
running on an IBM compatible computer with a 14 inch monitor
and sound system. This version used blocks as compared to
disks or rings used in a traditional Tower of Hanoi puzzle.
The blocks were numbered and colored to facilitate the
experiment. Figure 1 is a picture of the screen display
from the computer program.
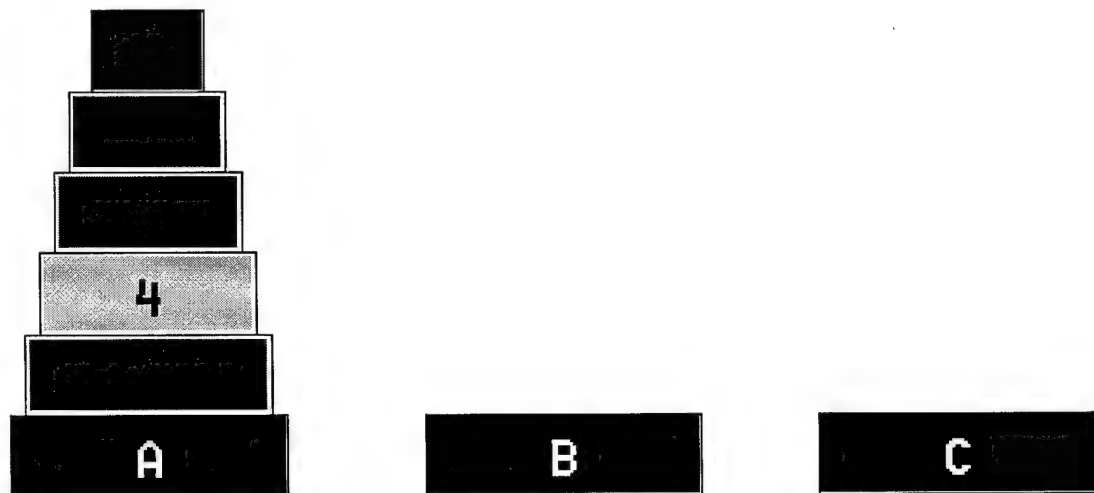
Figure 1. Computer version of Tower of Hanoi.

Procedure. Female and male participants were randomly assigned to either the relevant or irrelevant practice conditions. Results of previous pilot work suggested that the 4-disk version of the Tower of Hanoi (as compared to a 3-disk version) would optimally affect the performance on the computerized 5-disk version. Therefore participants in the relevant practice condition solved the 4-disk Tower of Hanoi on the wooden model three times with no specific time limit. Participants in the irrelevant practice condition worked on the dinosaur version of the Eight Puzzle for ten minutes. The instructions to the participants are in Appendix C.

Following the practice, all participants completed the 5-disk computer version of the Tower of Hanoi. Instructions were presented both visually and auditorially on the computer (see Appendix D). All moves were made by the experimenter in response to the participant's instructions such as: "move block 1 to A" or "move block 1 to block 2." The moves and response times were directly recorded in a computer data file.

Results and Discussion

The total number of moves and times were calculated for each participant and the average number of moves and solution times for the four conditions are presented in

17

Table 1. A 2 X 2 (gender X practice) analysis of variance was used for both the move and time data. The ANOVA indicated a main effect of type of practice for number of moves, $F(1,39) = 8.12$, $p = .01$ and total time, $F(1, 39) = 6.67$, $p = .01$. Specifically, participants who received relevant practice completed the computer task in less moves and the moves were made faster than for participants who received the irrelevant practice. There were no significant interactions of gender and practice for any of the outcome measures. However, there was a main effect of gender for total time and time per move indicating faster performance for males relative to females, $F(1,39) = 5.27$, $p = .03$ and $F(1,39) = 3.39$, $p = .04$, respectively.

Table 1

Mean (standard deviations) of Number of Moves, Total Time, and Time Per Move by Gender and Practice Conditions

| | Gender | | | | | | | | |
| | Female | | | Male | | | Total | | |
Practice	Moves	Total Time	Avg Time	Moves	Time	Avg Time	Moves	Total Time	Avg Time
Relevant	49 (10)	257 (73)	5.2 (1.0)	52 (13)	240 (105)	4.5 (1.5)	51 (11)	248 (89)	4.9 (1.3)
Irrelevant	68 (21)	536 (371)	7.7 (4.8)	63 (18)	307 (156)	4.6 (1.2)	65 (20)	421 (301)	6.2 (3.8)
Total	59 (19)	396 (297)	6.5 (3.6)	57 (16)	273 (134)	4.5 (1.3)	58 (18)	335 (236)	5.5 (2.9)

The results support the hypothesis that practicing a 4-disk version of the Tower of Hanoi on the wooden model facilitates performance on the computer version of the Tower of Hanoi. Also as previously reported in the literature males solve the Tower of Hanoi faster than females. Following the confirmation of these results the main experiment was conducted.

Experiment 2

Purpose

This experiment was designed to investigate the effects of gender mix and prior practice on team decision making while solving the Tower of Hanoi. Team performance was measured on the computerized Tower of Hanoi in the same manner as for individuals in the previous experiment. Relevant and irrelevant practice was also manipulated by using the same materials used earlier, the wooden Tower of Hanoi and the Eight Puzzles. A between subjects design with two factors -- team mix (female-female, male-male, and female-male) and gender receiving relevant practice (female, male, or neither) was used in this experiment, resulting in seven treatment conditions. There were two combinations of a complete factorial combination of these variables that were not formed due to mutual exclusion -- males could not

20

receive practice in a female-female mix and females could not receive practice in a male-male mix.

Method

Participants. A total of 110 two-person (dyads) teams participated in this study (116 women and 104 men, mean age = 18). All participants were undergraduates in a general psychology class who participated as part of a course requirement. The results of only 84 teams (84 women and 84 men), 12 under each of the 7 treatment conditions were used for data analysis. Teams were eliminated from analysis based on the following criteria: (a) total moves over 100, (b) total time over 30 minutes, (c) high degree of familiarity between team members, (d) either or both participants were familiar with the Tower of Hanoi, or (e) equipment failure (computer or tape recorder). There was no evidence that teams were systematically eliminated from one experimental condition relative to the others. All volunteers read and agreed to the conditions in the attached (see Appendix A) informed consent approved by the Florida State University human subject committee.

Participants were randomly paired by gender to form three different groups: female-female (FF), female-male (FM), and male-male (MM). The teams participated in one of three conditions; (a) a male receiving relevant practice,

(b) a female receiving relevant practice, or (c) neither
participant receiving relevant practice. The combinations
of gender mix and gender receiving relevant practice
produced seven different experimental groups.

Procedure. When the participants arrived at the
laboratory they were directed to two separate rooms, briefed
on the purpose of the experiment, and told that their
session would be audio-recorded. The same instructions
(Appendix C) used in Experiment 1 were read to the
participants and they practiced their assigned task for 10
minutes. Depending on their assigned experimental
condition, participants practiced on the wooden Tower of
Hanoi (4-disk) or one of two Eight Puzzles. The only data
collected during this phase of the experiment were the
number of times participants completed the puzzles.
Participants completed a written retrospective protocol
following their practice task (see Appendix E) and then
proceeded to the team phase of the experiment.

The teams involved individuals seated next to each other
in front of the computer and each participant wore a lapel
microphone connected to the tape recorder. As in the
validation study, instructions were presented visually and
over the speakers with some additional instructions
concerning team procedures (see Appendix D). Participants

22

were required to alternate which team member gave the move instructions and prefaced all move instructions with "Jack move." This assured that there was no confusion between the participants and the researcher (Jack) when teams were discussing moves as opposed to actually making a move. All moves, time for each move, and errors (large block on smaller block or moving a block not on top of stack) were recorded on the computer.

Once the teams completed the task they were asked to provide a written retrospective protocol. The instructions for the protocol were the same as used in the first phase of the experiment (Appendix E) except they were given 5 minutes to complete the task and on top of the form there was a Likert scale from 1 (just met) to 5 (best friends) to determine team familiarity. Participants were thanked for their participation and provided with a summary describing the research.

Data collection. As previously mentioned, all moves, time for each move, and errors were recorded on the computer. All sessions were recorded on a stereo tape recorder with one participant's verbalizations recorded on the left channel and the other on the right channel. These stereo recordings helped facilitate the written transcriptions, reducing errors resulting from confusing the

23

speakers. The written transcriptions preserved speaker identification, so that the data from the practice phase could be associated with the correct team member. Segments on the written transcripts were delineated by a change of speaker. A coding sheet was used to code all transcribed verbalizations using the Multiple Protocol Analysis System (MPAS) computer program (Cruthcer, 1994). This program randomly presents transcription segments for coding. Coding of the transcription data was used to determine if gender mix or variations in practice affected decision making processes, outcomes, or interactions between team members.

Results

Several different analysis were conducted on the various forms of data collected during this experiment. The first section describes the results from the objective performance data collected on the computer. This includes the total number of moves, total time, and time per move for teams solving the computerized Tower of Hanoi. The teams' moves were then plotted on a diagram containing all possible moves for the 5-disk Tower of Hanoi and indicators (reversals, circularity, persistence down wrong path, and minimum path) from these data were analyzed. Next, the communications between team members were analyzed from the written transcriptions. The final section contains the

24

results from the retrospective protocols participants completed following the computerized Tower of Hanoi task.

Multiple regression procedures are used in most of the following analyses. Regression analysis was used instead of analysis of variance because the nature of the design resulted in an incomplete factorial combination of gender mix and relevant practice gender. Gender mix and relevant practice gender factors were transformed into more interpretable contrast variables based on theoretical considerations. The contrasts are not completely orthogonal because of the incomplete design, but the correlations are small and will not be considered a serious threat to the statistical analysis. These contrasts are: (a) homogenous versus heterogeneous teams, (b) female-female versus male-male teams, (c) teams with one member receiving relevant practice versus no-practice, and (d) females receiving relevant practice versus males receiving relevant practice. The overall significance for each regression model will be reported with the statistical tests for the contrasts presented in a table. None of the interactions between the contrast variables were significant and therefore are not reported. This procedure uses protected testing to guard against the potential inflation of error rate due to

multiple statistical comparison tests. All t-tests are two-way tests unless stated otherwise.

Performance data. The total number of moves and times were calculated for each team from the computer collected data. The average number of moves and standard deviations for all conditions are presented in Table 2. Table 3 profiles summary of the total time and standard deviations to complete the computerized version of the Tower of Hanoi. The average times per move and standard deviations are presented in Table 4.

Table 2

Mean (standard deviations) Number of Total Moves by Gender Mix and Relevant Practice Gender

| | Relevant Practice Gender | | | |
Gender Mix	Female	Male	Neither	Total
Female-Female	62.3(17.2)		63.3(19.5)	62.8(18.0)
Male-Male		46.3(10.9)	53.9 18.2)	50.1(15.2)
Female-Male	63.4(18.5)	46.4(14.5)	53.1(12.5)	54.3(16.5)
Total	62.9(17.5)	46.3(12.6)	56.8(17.2)	55.5(17.1)

Table 3

Mean (standard deviations) of Total Times in Seconds by
Gender Mix and Relevant Practice Gender

| | Relevant Practice Gender | | | |
Gender Mix	Female	Male	Neither	Total
Female-Female	678(378)		679(322)	678(343)
Male-Male		328(128)	550(313)	439(260)
Female-Male	639(329)	443(245)	762(322)	615(321)
Total	659(347)	385(200)	664(322)	583(323)

Table 4

Mean (and standard deviations) of Average Time Per Move In
Seconds by Gender Mix and Relevant Practice Gender

| | Relevant Practice Gender | | | |
Gender Mix	Female	Male	Neither	Total
Female-Female	10.3(4.2)		10.7(4.5)	10.5(4.3)
Male-Male		7.2(2.9)	10.1(4.1)	8.7(4.1)
Female-Male	9.8(3.8)	9.8(6.2)	14.3(4.5)	11.3(5.2)
Total	10.1(3.9)	8.5(4.9)	11.7(4.8)	10.3(4.7)

A graphic representation of the group means for total
moves to solution is presented in Figure 2. Viewing the
data on the graph it appears that there is a main effect of

relevant practice gender with males better (less moves) than
neither and females. It also appears that performance is
poorest in female-female teams, best for male-male teams. A
possible interaction of gender mix and which member receives
relevant practice is revealed in the finding that relevant
practice for one member of a female-female team has little
effect on performance while relevant practice for the
members of male-male teams results in a reduction in moves
to solution.
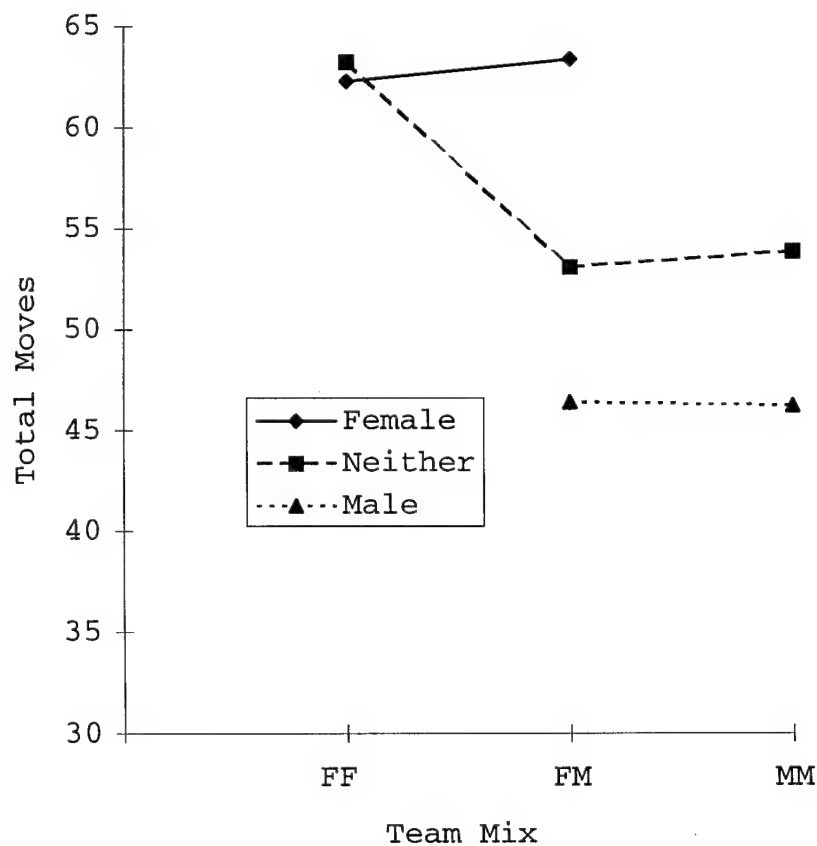
Figure 2. Mean number of moves by gender receiving relevant
practice and gender mix.

28

A standard multiple regression analysis was performed between total number of moves as the dependent variable and gender mix contrasts and relevant practice gender contrasts as independent variables. Based on theoretical considerations and inspection of the moves histogram, logarithmic transformations were used to normalize the distribution. Since the statistical outcomes were identical for both the transformed and untransformed data, only the results of the untransformed data will be presented.

The model R^2 of 0.17, reflecting the overall strength of relationship between number of moves and the independent variables, was statistically significant, $F(6,77) = 2.63$, $p = .02$. The adjusted R^2, compensating for the positive bias in R^2, was 0.11 and the standard error of estimate was 16.20. The effects of the individual contrasts are summarized in Table 5. The only significant contrast indicates that teams with males receiving relevant practice completed the puzzle in fewer moves as compared to teams with females receiving relevant practice.

Table 5

Regression Results Summary Table With Total Number of Moves

as the Dependent Variable ($\underline{N}$ = 84)

Contrast	$\underline{B}$	$\underline{SE}\ \underline{B}$	β	$\underline{t}$
Gender Mix				
Homogenous - Heterogeneous	0.26	0.52	0.06	0.55
MM - FF	-1.74	3.02	-0.08	-0.58
Relevant Practice Gender				
Males - Females	-8.50	3.31	-0.38	-2.57*
Relevant - Irrelevant	-0.24	0.51	-0.05	-0.46

*$\underline{p}$ < .05

As a second approach to understanding these data, two
separate 2 x 2 ANOVAs were calculated. One analysis was
computed on data from female-female and female-male gender
mix where either female members or no member of the team
received relevant practice. The second analysis was
computed for male-male and female-male gender mix teams
where either male members or no members received relevant
practice. The first analysis indicated no significant
interaction or main effects. The second analysis showed a
nearly significant gender receiving relevant practice
effect, $\underline{F}$ (1, 47)= 3.01, $\underline{p}$ = .09. Teams with males
receiving relevant practice completed the Tower of Hanoi in

less moves than teams with no members receiving relevant practice. In short these two analyses provided corroboration of the statistical results obtained in the overall regression analysis.

The means for the total time to completion data are illustrated in Figure 3. Inspection of the figure suggests a main effect of gender receiving relevant practice where teams with males receiving relevant practice take less overall time to solve the puzzle than the other two relevant practice factors. There also appears to be a trend indicating that female-female mixed teams are the slowest overall, followed by female-male, and then male-male, except when no member receives relevant practice the female-male group is the slowest.

830
730
630
530
430
330
230
130
30

Total Time

Female
Neither
Male

FF   FM   MM

Team Mix

Figure 3. Mean total time to completion by gender receiving
relevant practice and gender mix.

The model R^2 of 0.19, reflecting the overall

strength of relationship between number of moves and the

independent variables, was statistically significant,

$F(6,77) = 3.09$, $p = .01$. The adjusted R^2, compensating for

the positive bias in R^2, was 0.13 and the standard error of

estimate was 300.64. The effects of the individual

contrasts are summarized in Table 6. The only significant

32

contrast suggest a practice effect on overall time to complete the puzzle. Teams with one member receiving relevant practice completed the puzzle faster than teams without any practice.

Table 6

Regression Results Summary Table With Total Time as the Dependent Variable ($\underline{N}$ = 84)

Contrast	$\underline{B}$	$\underline{SE}$ $\underline{B}$	β	$\underline{t}$
Gender Mix				
Homogenous - Heterogeneous	-12.14	9.62	-0.13	-1.26
MM - FF	-71.59	56.13	-0.17	-1.28
Relevant Practice Gender				
Males - Females	-98.25	61.37	-0.23	-1.60
Relevant - Irrelevant	-22.64	9.62	-0.24	-2.35[*]

[*]$\underline{p}$ < .05

The average time per move is displayed graphically in Figure 4. The overall pattern is similar to the total time data except the male receiving relevant practice and female receiving relevant practice conditions do not show any strong differences. Both groups appear to be faster at each move than the group where no member receives relevant practice.
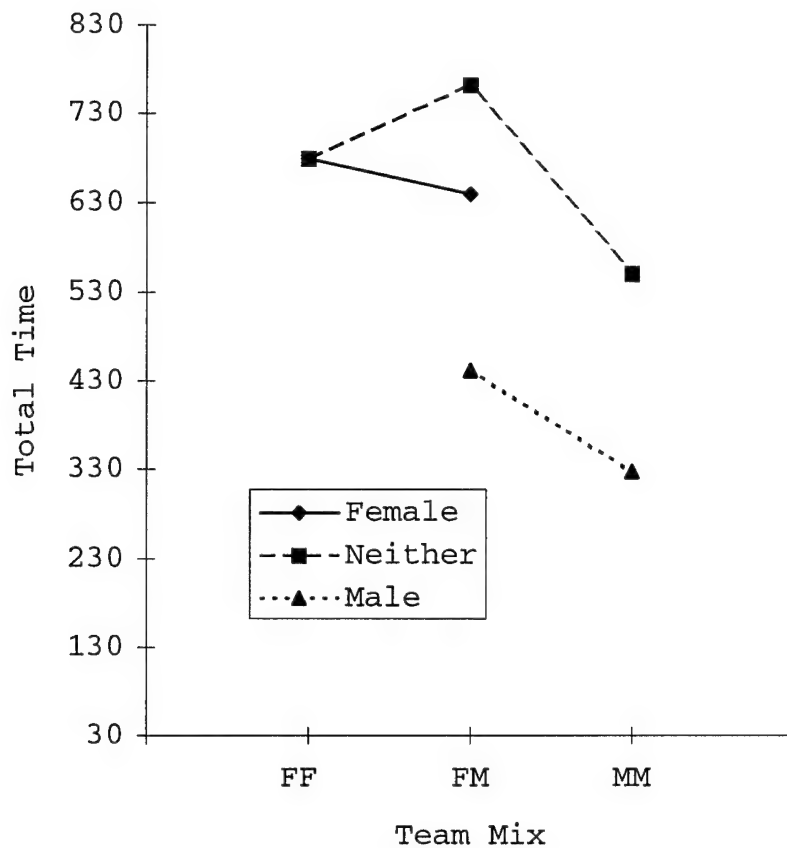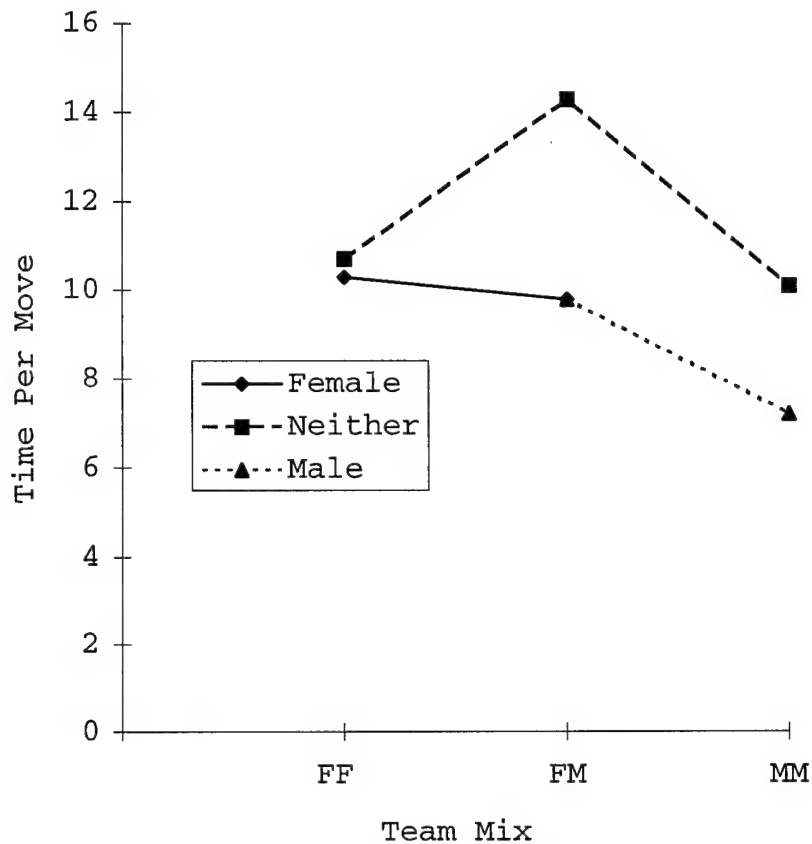
Figure 4. Mean time per move by gender receiving relevant
practice and gender mix.

The model R^2 of 0.17, reflecting the overall strength
of relationship between number of moves and the independent
variables, was statistically significant, $F(6,77) = 2.55$, p
= .03. The adjusted R^2, compensating for the positive bias
in R^2, was 0.10 and the standard error of estimate was 4.50.
The effects of the individual contrasts are summarized in
Table 7. Teams who had one member receive relevant practice
made faster moves as compared to teams where no one received

relevant practice. Also homogenous teams had shorter average move times as compared to heterogeneous teams.

Table 7

Regression Results Summary Table With Average Time Per Move as the Dependent Variable ($\underline{N}$ = 84)

Contrast	$\underline{B}$	$\underline{SE}$ $\underline{B}$	β	$\underline{t}$
Gender Mix				
Homogenous - Heterogeneous	-0.32	0.14	-0.23	-2.22[*]
MM - FF	-1.02	0.84	-0.16	-1.22
Relevant Practice Gender				
Males - Females	-0.01	0.92	-0.00	-0.01
Relevant - Irrelevant	-0.41	0.14	-0.30	-2.82[**]

[*]$\underline{p}$ < .05, [**]$\underline{p}$ < .01

Search Space Analysis. In addition to the analysis of the three performance measures discussed above, a finer grain analysis of the move patterns was conducted. This analysis was based on tracking the particular sequence of "moves" during problem solving.

There is a clear goal for this task and the minimum number of moves is 2^n-1 where n = the number of disks. Therefore in a 5-disk problem 31 is the minimum number of moves. Simon (1975) identified several strategies which

35

could be used to solve the puzzle -- goal-recursion strategy, perceptual strategy, sophisticated perceptual strategy, and a move-pattern strategy. Depending on which strategy is used there will be different demands placed on short-term memory (STM), different sets of concepts required, and different perceptual tests made for the execution of the strategies. For example, if a participant uses the goal-recursion strategy they reduce the problem from the overall goal (move the stack from peg A to peg C) to subgoals (move a single disk) until they can make a legal move. This strategy requires a way of representing goals internally, and holding them in STM while completing the subgoals. Therefore the information in STM is necessary and sufficient for the goal-recursion strategy. In contrast to Simon's research, which was based on individuals solving the puzzle, in the present study the results of a search space analysis was used to determine strategy differences for team solutions.

Hidden within the dependent measure of total number of moves may be other indices that differentiate teams based on which gender receives relevant practice and gender mix. Figure 5 illustrates all possible moves for the 5-disk version of the Tower of Hanoi. The top of the triangle represents the start state with all five disks stacked on

36

peg A (a54321bc) and the lower right corner of the triangle represents the goal state with all disks stacked on peg C (ac54321b). The quickest move pattern (31 moves) is from the top of the triangle down the right side to the bottom right corner. The moves for each team were plotted on the search space diagram and several component measures were computed from the diagrams. Figure 6 is an expansion of the top portion of the search space diagram with examples of the various measures derived from the plotted team moves.

All Disks on Peg A

Quickest Path

All Disks on Peg B

All Disks on Peg C

Figure 5.   Search space for 5-disk Tower of Hanoi (quickest path in bold, 31 moves).

38
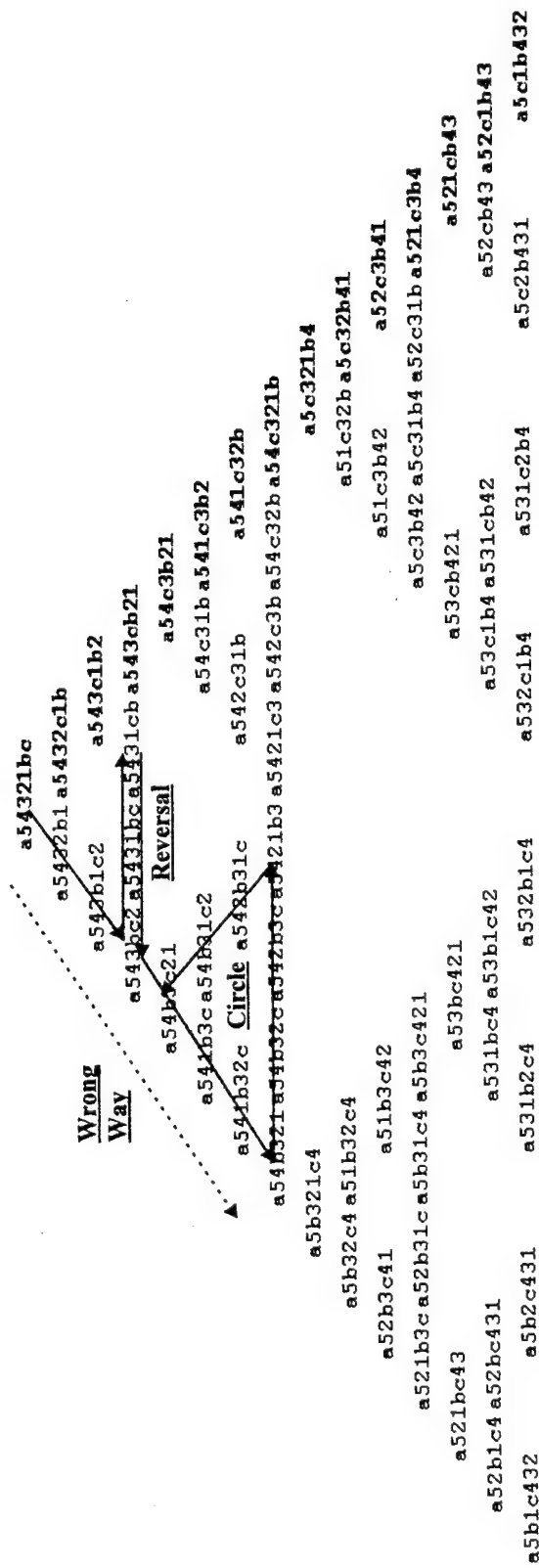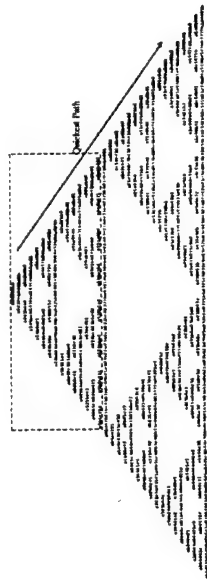
Figure 6. Search space blow-up illustrating reversals, circularity, and persistence down the wrong path (wrong way).

Four different measures were computed from the paths of individual teams on the search space diagrams: (a) circularity, (b) reversals, (c) wrong way, and (d) minimum path. Circularity refers to a pattern of moves that started and returned to the same space via a different path. This measure yielded a total number of moves within each circle for a team. Reversals were paths that retraced previous moves and could be any number of moves. These two indices were measured on a continuous scale, the last variable, persistence down the wrong path (wrong way), is really a categorical variable. A number between 1 and 7 was assigned to a team depending on how far down the left side of the triangle their path traveled. This variable approximates a continuous variable and is treated as such in all the analyses. Finally, an overall minimal path variable was calculated by subtracting moves classified as reversals and circularity from the total number of moves. This variable is defined as the minimum path the team would have taken devoid of any returns to the same point. The results are presented in Table 8.

Table 8

Mean (standard deviations) of Minimum Path, Reversals,

Circles, and Wrong Way by Gender Mix and Relevant Practice

Gender

	Relevant Practice Gender			
Gender Mix	Female	Male	Neither	Total
Female-Female				
Min Path	57.8(15.8)		55.9(15.4)	56.9(15.3)
Reversals	3.3(4.1)		5.0(4.4)	4.1(4.3)
Circles	1.3(4.3)		2.3(5.5)	1.8(4.9)
Wrong Way	2.4(2.0)		1.8(2.3)	2.1(2.2)
Male-Male				
Min Path		45.2(11.0)	48.3(13.1)	46.8(12.0)
Reversals		1.1(1.5)	3.1(3.2)	2.1(2.7)
Circles		0	2.5(5.3)	1.3(3.9)
Wrong Way		1.3(2.1)	1.4(2.15)	1.3(2.1)
Female-Male				
Min Path	59.2(17.3)	45.3(13.5)	50.7(11.2)	51.8(15.0)
Reversals	4.3(5.5)	1.1(1.3)	2.4(2.6)	2.6(3.7)
Circles	0	0	0	0
Wrong Way	2.6(2.2)	0.8(1.6)	1.2(1.6)	1.5(1.9)
Total				
Min Path	58.5(16.2)	45.3(12.1)	51.6(13.3)	51.8(14.6)
Reversals	3.8(4.8)	1.1(1.4)	3.5(3.6)	2.9(3.7)
Circles	0.6(3.1)	0	1.6(4.5)	0.9(3.4)
Wrong Way	2.5(2.0)	1.0(1.9)	1.4(2.0)	1.6(2.0)

The model R^2 of 0.17, reflecting the overall strength of relationship between minimum path and the independent variables, was marginally significant, $F(6,77) = 2.08$, $p = .06$. The adjusted R^2, compensating for the positive bias in R^2, was 0.07 and the standard error of estimate was 14.08. The effects of the individual contrasts are summarized in Table 9. The only significant contrast indicates that teams with males receiving relevant practice had a shorter minimum path compared to teams with females receiving relevant practice. These results do not yield any new information concerning group differences and next we look at more detailed indicators of the team's moves patterns.

Table 9

Regression Results Summary Table With Minimum Path as the Dependent Variable (N = 84)

Contrast	B	$SE\ B$	β	t
Gender Mix				
Homogenous - Heterogeneous	0.03	0.45	0.01	0.06
MM - FF	-1.29	2.63	-0.07	-0.49
Relevant Practice Gender				
Males - Females	-6.92	2.87	-0.36	-2.41*
Relevant - Irrelevant	0.05	0.45	0.01	0.10

*p < .05

42

The model R^2 of 0.14, reflecting the overall strength of relationship between reversals and the independent variables, was marginally significant, $F(6,77) = 2.12$, $p = .06$. The adjusted R^2, compensating for the positive bias in R^2, was 0.07 and the standard error of estimate was 3.53. The effects of the individual contrasts are summarized in Table 10. The only significant contrast indicates that teams with females receiving relevant practice made more reversals compared to teams with males receiving relevant practice. Even though these results are only marginally significant they provide a reasonable explanation for the relevant practice gender effects on total moves. The reason teams with females that receive relevant practice make more moves is because they make more reversals along their move paths.

Table 10

Regression Results Summary Table With Reversals as the

Dependent Variable ($\underline{N}$ = 84)

Contrast	$\underline{B}$	$\underline{SE}$ $\underline{B}$	β	$\underline{t}$
Gender Mix				
Homogenous - Heterogeneous	0.06	0.11	0.06	0.52
MM - FF	-0.13	0.66	-0.03	-0.19
Relevant Practice Gender				
Males - Females	-1.58	0.72	-0.33	-2.20[*]
Relevant - Irrelevant	-0.14	0.11	-0.13	-1.22

[*]$\underline{p}$ < .05

So far the results have focused on the data concerning the teams' moves and time for each move. These data help to explain the outcome behavior (what the teams accomplished) while the next section reviews data from the team communications.

Communication analysis. All sessions were recorded on a stereo tape recorder with one subject's verbalizations recorded on the left channel and the other on the right channel. This method was used to facilitate accurate transcriptions, insuring that the verbal segments were attributed to the correct person talking. All move instructions were deleted from the transcriptions since

these data were analyzed previously. The transcriptions were then divided into segments according to changes in speaker (see Appendix F for an example transcript). Before discussing these data it is important to discuss how these data were analyzed using the Multiple Protocol Analysis System (MPAS) computer program (Cruthcer, 1994).

The MPAS program implements a localized encoding scheme enabling the independent coding of segments. Segments are presented randomly from the entire transcription and are then coded without knowledge of prior or post segments. Two different coders can code the data and the results can be compared, computing interrater reliabilities.

The coding scheme used for this analysis was adapted from a study attempting to understand the communication process between crewmembers in a flight simulator mission (Kanki and Foushee, 1989). The following coding categories were used: (a) commands, (b) explanations, (c) suggestions, (d) acknowledgment, (e) planning, (f) question, (g) utterance, and (h) other. The coding was restricted so that each segment was encoded using only a single category. Commands mainly consisted of statements where one person told the other what move to make. Participants used explanations to justify a move or clarify rules/strategy to the other member. Suggestions were similar to commands, but

45

not as forceful -- "I think you should move this next."
There were a variety of questions, but the most common dealt
with one person asking the other for input on the next move.
Planning statements were the longest segments and consisted
of any segment that discussed more than two moves ahead.
Acknowledgments were simple yes, no, or OK responses.
Utterances were segments such as um, ha, and ah. All other
segments were coded as other. See Appendix G for several
examples of how statements from the written transcripts were
coded.

The written transcripts were coded by one researcher
for all teams and a random sample of 10 transcripts were
coded by a second researcher to determine the interrater
reliability. The interrater reliability was 82% using a
point-by-point agreement method. The means for total number
of verbal segments are presented in Figure 7 for the
combinations of gender mix and relevant practice gender. A
regression analysis did not yield any significant results
among groups. An inspection of Table 11 suggests that the
lack statistical significance is most likely a result of the
high variability in the data. Another approach to
understanding the verbal data involves the expected
communication patterns based on which person received
relevant practice.
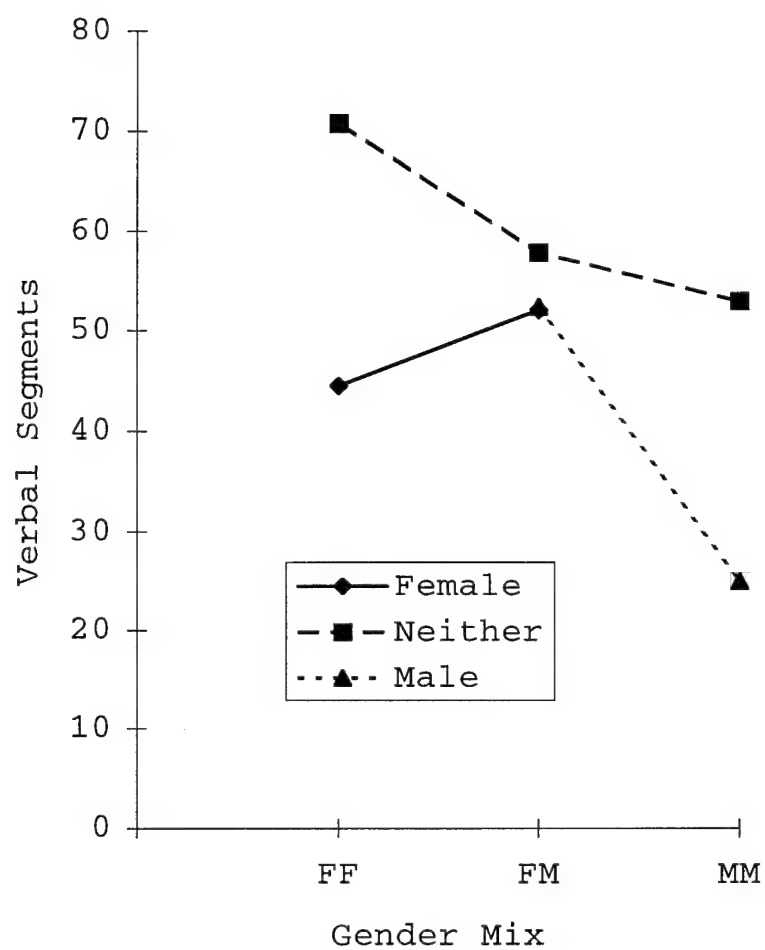
Figure 7. Mean number of verbal segments by gender receiving relevant practice and gender mix.

Table 11

Mean (standard deviations) of Total Number of Verbal

Segments by Gender Mix and Relevant Practice Gender

| | Relevant Practice Gender | | | |
Gender Mix	Female	Male	Neither	Total
Female-Female	44.5(37.6)		70.8(42.6)	57.7(41.5)
Male-Male		25.0(17.6)	53.0(51.0)	39.0(40.0)
Female-Male	52.1(29.9)	52.4(43.1)	57.9(48.2)	54.1(40.0)
Total	48.3(33.4)	38.7(35.1)	60.6(46.7)	50.8(40.7)

Considering only those groups where one team member
received relevant practice there should be differences in
the verbal categories based on who received the relevant
practice. The person with the most knowledge of the
situation should produce more commands, planning,
suggestions, and explanations than the "novice." Likewise
the person who received the irrelevant practice should ask
more questions than the "expert." Figure 8 shows the
differences for each verbal category by type of practice in
each of the relevant practice conditions.
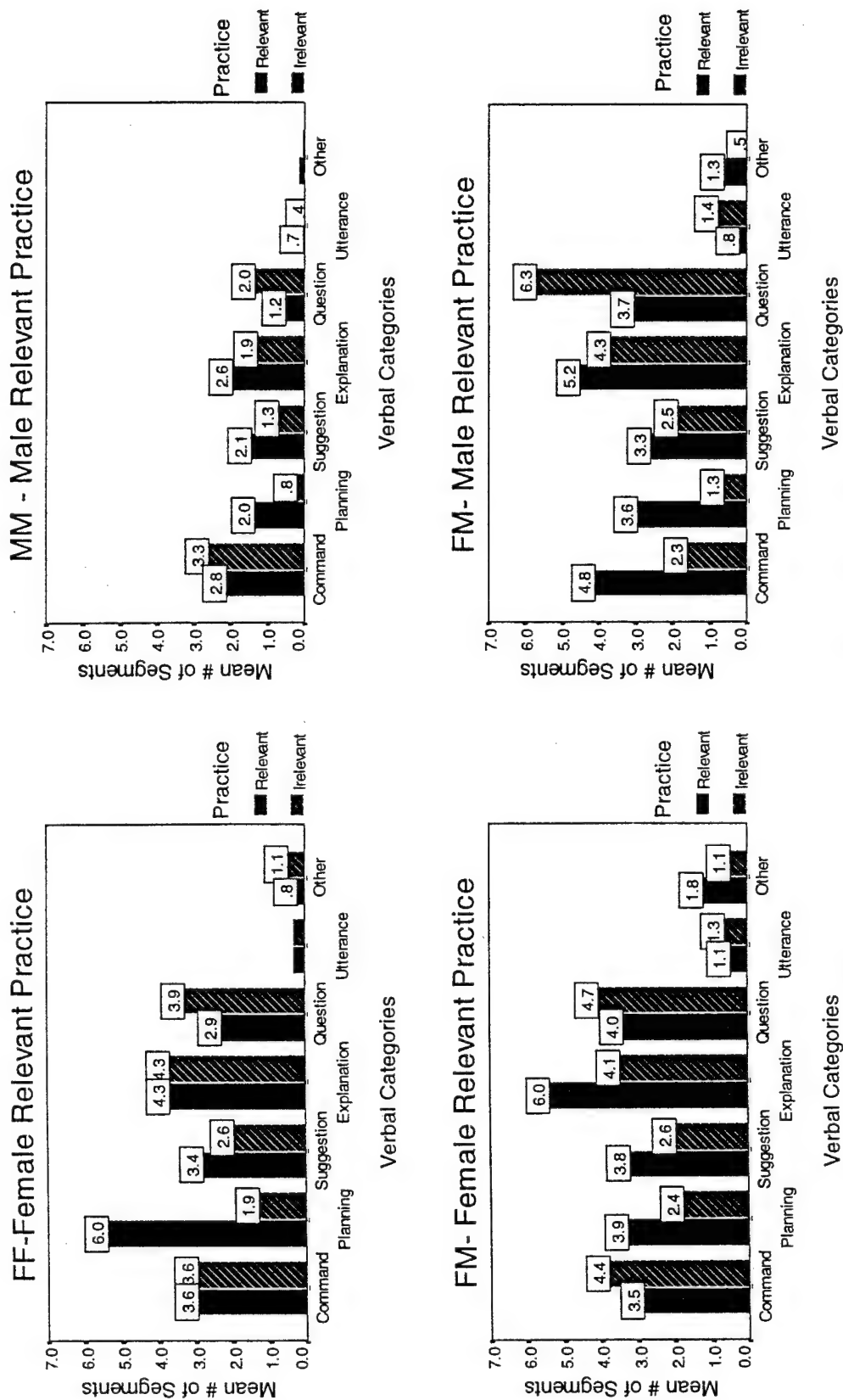
Figure 8. Mean number of verbal segments by gender mix and practice type.

A visual inspection of Figure 8 reveals some trends that are counter to the expected pattern of results. The most interesting trends involve the number of commands for each group. Focusing on the same sex groups represented by the top two bar graphs, the female-female group showed no differences between the person receiving relevant practice and the other person for number of commands. The trend in the male-male group shows that the person receiving relevant practice gave less commands than the other person. There appears to be a greater difference in planning behaviors between the person receiving practice and the other individual for the female-female group when compared to the male-male group.

It was expected that commands would be the best indicator to help explain group differences according to the verbal data. Therefore statistical analysis was only computed for the command category data. The results for the command category in the female-male group with male receiving relevant practice showed a significant difference, $t(22) = -1.82$, $p = .04$, one-tailed, when females received relevant training males give more commands. Even in the mixed gender group where both males and females receive irrelevant practice visual inspection of Figure 9 suggests a trend where males are giving more commands than females.

The results of the data for all female-male groups combined (with or without team member receiving relevant practice) show a significant effect with males making more commands than females, $\underline{t}(22) = -2.16$, $\underline{p} = .03$, one-tailed.

FM- No Relevant Practice
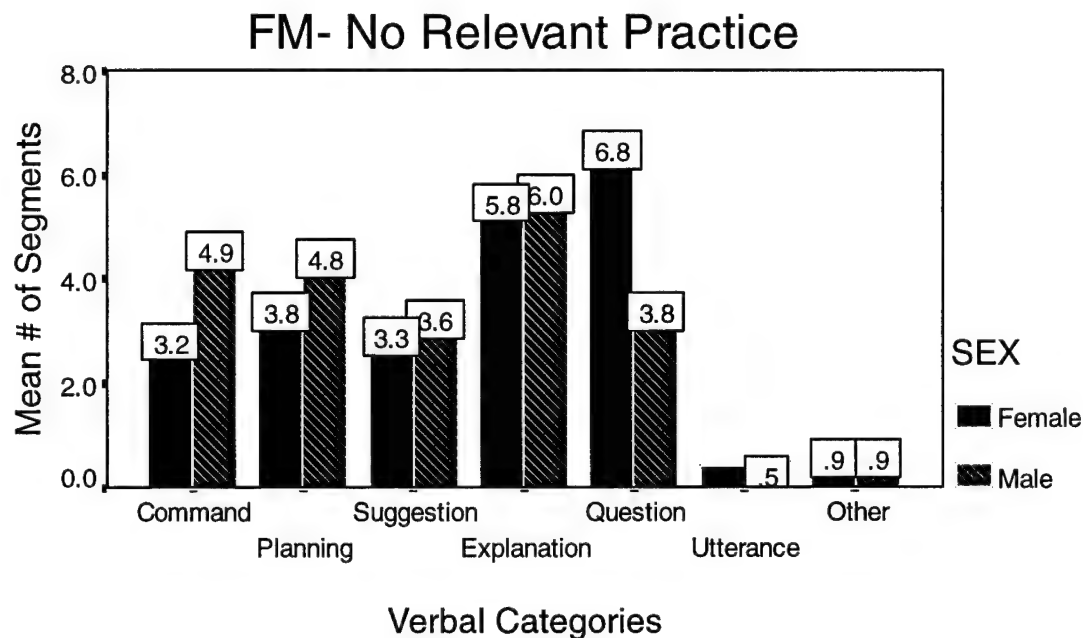
Figure 9. Comparison of verbal segments in female-male, no relevant practice group.

It was expected that some of the participants who received relevant practice would convey this fact to their teammate at some point during the session and that this would impact the team's outcome. The communication data revealed that only 8 out of a possible 48 participants actually made a statement about their experience with the wooden Tower of Hanoi. Three of the participants were in

51

the female-female group, two in the female-male with female
receiving relevant practice, two in the female-male with
male receiving relevant practice, and one in the male-male
group. In short, since the participants revealing their
"expertise" were spread almost evenly amongst the groups,
there were no significant results from these data.

Retrospective verbal reports. Data were collected from
both subjects following the team session, concerning what
they were thinking while they were solving the puzzle. The
hand-written reports were not very long, averaging only a
half page. A cursory content analysis indicated that the
most frequent comments all groups combined were: (a) thought
problem was easy at first, (b) worked well with partner, (c)
felt stupid, (d) confused, (e) frustrated, (f) acted as
leader, (g) thought partner was smarter than I, (h) would
have been quicker working alone, and (i) tried to solve
problem during instructions. There was some other
insightful information concerning the experimental design,
but no further analysis was accomplished since there was not
a sufficient amount of data.

Discussion

The results from the first experiment show no gender differences regarding number of moves to complete the Tower of Hanoi puzzle when individuals worked alone. However, the data from the first experiment did show a gender difference in terms of time to solve the puzzle, a finding consistent with prior research. This holds true for subjects regardless of prior experience with the Tower of Hanoi. The results also show that regardless of gender, prior practice on a wooden 4-disk Tower of Hanoi will facilitate performance on an analogous 5-disk computer version.

Some of the following refers to results that are "marginally significant" or trends. I will not devote much time to the issue of whether one should or should not mention these findings, but the issue of statistical power concerning the current study needs to be addressed. A conservative estimate of the power for the regression analysis with moves as the dependent variable from Experiment 2 is .26 for the gender mix factor (Cohen & Cohen, 1983). Therefore the probability of failing to reject the null hypothesis when it is false (Type II error) is .74 (1 - power). According to Cohen and Cohen this is considered low power and one should consider ways to increase power to an acceptable level around .80. This

could be accomplished by increasing the number of subjects or decreasing variability. Based on the information concerning the statistical power in this study I will discuss trends and also address some changes that may reduce the variability.

The team mind model (Thordsen and Klein, 1989) provides a framework for discussing the results of Experiment 2 in which teams attempted to solve the Tower of Hanoi problem. The behavioral level, which refers to the team's overt actions, was assessed by collecting information about each move on the computer. The collective consciousness was reflected in the verbal transcripts collected during the task. Finally, the subconsciousness level, which refers to individual knowledge that is not shared with the other team member, was assessed by two different methods. First, knowledge concerning the Tower of Hanoi was manipulated experimentally. Some participants received practice on the Tower of Hanoi and others did not. Therefore individual knowledge was controlled for, a priori. Also written retrospective verbal reports were collected, providing further insight into the subconsciousness level.

The results from the current study suggest that gender is a factor when teams are solving the Tower of Hanoi puzzle. Specifically when a male on a team has prior

54

practice with the puzzle the team will solve the more difficult computerized version in fewer moves and in a shorter overall time as compared to when the female team member has prior practice. Also, considering only female-female teams, the teams appear to show no improvement when one individual receives practice compared to no practice for either individual.

While these data show that problem solving behavior is not improved when females receive relevant practice versus males, there is no indication that this finding would generalize to all situations. There are many other factors such as stress, expertise, and maturity that were not examined in the current study. I do think this study represents a reasonable methodology which can be used in more naturalistic studies. Why practiced males performed better will be addressed later from the results of the search space and the verbal data.

The conclusions from the time per move data are consistent with data from the expertise literature which might suggest that an expert at chess would make quicker moves than a novice. Similarly the results from the current study show that that when a participant received practice, regardless of gender, the team's average time per move was shorter than when nobody on the team received prior

practice. The other result concerning average move times
suggests that same gender teams make faster decisions (not
necessarily better) concerning moves. Perhaps males and
females actually make decisions differently and when forced
to make team decisions in mixed teams it interferes with
their usual decision making processes.

The results of the collective consciousness measures
reveal some subtle group differences concerning how team
members communicate with each other. It appears that when
males are on a team, regardless of prior experience, they
will make more statements telling the other person what to
do (commands) as compared to females. This could be a
partial explanation for the gender effect on number of moves
and total time found in this experiment. This finding has
many implications for team decision making in real-world
scenarios if it remains invariant across situations. An
important goal of team decision making research should be to
determine the effects and then understand how to eliminate
any detrimental effects. Assuming males always make more
commands, training programs could address this fact and how
teams should best cope with it.

Another account of the gender effect is based on the
search space results. Teams with females who have practiced
the Tower of Hanoi tend to reverse moves more often than

teams with males who had practice. Knowing this fact it may be possible to improve team performance by designing training which incorporates information concerning reversals. This result also hints at the possibility that there was a differential effect of practice on the participants. Perhaps a modification to a future study would include more data collection during the practice phase and maybe change this phase from practice to a "train to criterion" approach. This approach would require all individuals to solve the Tower of Hanoi until they are able to do it in some set number of moves. This would afford greater control over what effect practice has on the participants.

The prior individual experience concerning the tower of Hanoi was controlled by the research design. Except for participants who had extensive experience with some form of the Tower of Hanoi before the experiment, and were therefore eliminated from the experiment, all participants were exposed to the same practice conditions. The low frequency of sharing facts (subconsciousness level) about one's knowledge concerning the Tower of Hanoi is consistent with many aviation accident reports indicating that a crewmember had critical information that was not shared with the rest of the team.

57

Previous research with aircrews (Foushee et al., 1986; Woody et al., 1994; Barker et al., in press) suggests another factor that can impact team performance. As the level of familiarity between the pilots increases, at some hypothetical point team performance declines. The current study avoided the possibility of a familiarity confound by eliminating data from teams that rated themselves as highly familiar with each other. This factor should be incorporated in future research and the level of familiarity between team members could be investigated by using roommates, friends, or co-workers as participants. This would allow us to determine how familiarity interacts with gender (if it does) during decision making.

Another dimension that warrants further exploration is the effect of expertise on team decision making. The current study utilizes naive participants rather than individuals with a high level of domain area expertise. Would the pattern of results differ for individuals/teams considered to be expert problem solvers? In many real-world situations the expertise gradient between team members would be much greater than in the current study. Other factors such as standard operating procedures, company policies, and social interactions complicate team decision making. Additionally, research conducted on cooperative learning

58

suggests that training the individuals to work as a team may improve team performance. The participants in the current study had no practice working in teams prior to starting the computerized version of the Tower of Hanoi. This poses several questions worth investigating.

Earlier I alluded to the ecological validity of this study. This research was designed primarily to develop a reasonable methodology that can be used to study team decision making both in the laboratory and in more naturalistic situations. The results suggest that (for undergraduates at least) there are differences in problem solving that are influenced by gender, but I have not answered whether or not these differences remain following extensive training in a particular domain. I do present a methodology and model that should assist future projects that attempt to understand factors affecting team decision making.

APPENDIX A

INFORMED CONSENT FORM

I freely and voluntarily and without any element of force or coercion, consent to be a participant in the research project entitled "Team Decision Making".

It is being conducted by John M. Barker, Jr, M.S., who is a graduate student in the Psychology department at Florida State University. I understand the purpose of his research project is to better understand how teams make decisions. I understand that if I participate in this project my conversations will be recorded while working on the experiment with another student.

I understand my participation is totally voluntary and I may stop participation at any time. All my answers to the questions will be kept confidential and identified by a subject code number. My name will not appear on any of the results. No individual responses will be reported. Only group findings will be reported.

I understand that there are no risks involved, but I may stop my participation at any time I wish.

I understand there are benefits for participating in this research project. First, I will learn about conducting psychological research. Also I will help researchers better understand factors that impact team decision making. This may in turn help to improve areas where team make decisions such as airline crews and hospital surgery teams.

I understand that I may contact John M. Barker, Jr. at the Florida State University Psychology department, KRB-128, (904) 644-4382, for answers to questions about the project. Group results will be sent to me upon my request.

I understand that this consent may be withdrawn at any time without prejudice, penalty or loss of benefits to which I am otherwise entitled. I have been given the right to ask and have answered any inquiry concerning the study. Questions, if any, have been answered to my satisfaction. In the future, I understand I may contact John M. Barker, Jr., Florida State University Psychology Department, KRB-128, (904) 644-4382, for answers to questions about this research or my rights. I have read and understand this consent form.

APPENDIX B
PICTURES OF EIGHT PUZZLES USED IN EXPERIMENTS

APPENDIX C
INSTRUCTIONS TO PARTICIPANTS FOR EXPERIMENT 1

Instructions for Tower of Hanoi

For the next few minutes I would like you to work on
solving the puzzle in front of you. The goal is to
move the 4 disks from there current location to the far
right peg as quickly as possible but also in as few
moves as possible.
The rules are as follows: You can only move 1 disk at
a time and you can not place a larger disk on top of a
smaller disk. Any questions? Please begin.

Instructions for 8-puzzle

For the next few minutes I would like you to work on
solving the puzzle in front of you. The goal is to
move the blocks so that they recreate the picture in
front of you. Do this as quickly as possible but also
in as few moves as possible. Any questions?
Please begin.

INSTRUCTIONS PRESENTED ON THE COMPUTER FOR TOWER OF HANOI IN
EXPERIMENTS 1 AND 2 (**BOLD** EXPERIMENT 2 ONLY)

Screen 1

Welcome to "Stack the Blocks".
Your goal is to move the stack of blocks from Base A to Base
C as quickly as possible in the minimum number of moves.

Screen 2

There are three rules that you must follow while completing
this task.

Screen 3

1. You can only move one block at a time and it must be on
the top of a stack.

Screen 4

2. You can not place a larger block on top of a smaller
block.

Screen 5

3. The experimenter will make all the moves according to
your instructions.

For example if you want to move Block 1 to Base B, say "move
1 to B".

Screen 6

And Block 1 will be moved to Base B.

If you want Block 1 on top of Block 4, just say "move 1 to
4".

Screen 7

And Block 1 will be stacked on top of Block 4.
**This is a team exercise, so work together to solve this
problem**

Screen 8

<u>**WARNING!**</u>

**You must alternate between each team member giving move instructions to the experimenter.
Begin each move instruction with "JACK move" and then your move. This will avoid any confusion while you and your team member are discussing moves.**

Screen 9

Any Questions?

Remember that you are trying to complete the task as quickly as possible in the minimum number of moves.

Please let the experimenter know when you are ready to begin.

APPENDIX E
INSTRUCTIONS FOR RETROSPECTIVE PROTOCOL

Retrospective Step

Now I want to see how much you remember about what you were
thinking from the time that I gave you the task and when you
were finished. We are interested in what you actually can
remember rather than what you think you must have thought.
If possible I would like you to tell about your memories in
the sequence in which they occurred while working on the
problem. Start with your memories from when you first began
and work forward. Please tell me if you are uncertain about
any of your memories. I don't want you to re-work the
problem again, just report all that you can remember
thinking about when completing this form. Now for the next
3 minutes write down everything that you can remember.

EXAMPLE OF SEGMENTED TRANSCRIPT

ts59*
R: I think if we move 1,2, and then 1 and 3, then 4, then
stack 3 there. You can only move one at a time, right?\
L: Yeah\
R: 3B maybe?\
L: Yeah, sure. JM 3B\
R: JM 1A.   and if we put 2 there, 1 there, 4 there. I
don't know what we do after that\
R: JM 1B.   Then 4C?\
L: Yeah, JM 4C.   Put 1A, then....no\
R: If we move 1....let's move 14,25, no that wouldn't do
anything\
L: Oh, do 14, 25, 15\
L: JM 2B.   1A?\
R: JM 1C.   This one here...\
L: 5B, then 1A...no\
R: This one here, this one here, 2 there, 1 there, 3 there,
and move them around and get 4 there, 5 there.   So, JM
5B\
R: ooooh, 14,2 there\
L: Ok, your go\
R: JM 1A.   4B?\
L: JM 4B.   Then what?\
R: If we move 1,2,13\
L: Yeah, but you don't want 5 on the bottom\
R: Oh, you wanna invert it?\
L: I don't know, we gotta get 5C\
R: 5,4,3,2,1?\
L: Yeah
R: So, if we do 4 there, 1, we could put 2 here, 3 there,
move them over and put 5 there.   JM 4C\
R: JM 2B.   1B?\
L: No good, can't put...\
R: JM...no. Um, 3A and \
L: We are gonna get C completely cleared\
R: So 1,3,2,1,4. You think that will work?\
L: Um, well I'm trying to think. I have no clue\
R: If we put 3,1 then 1,2,3,4,5\
L: Then what do you do with the 4?\
R: Yeah

L: Um, I guess put...\
R: Put 1,2.   1,2,1, then 1,2,3,4; 5 there and shift those
there. Then we will be back to where we were\
L: Right and that won't do us good. If we could get 5C
then...\
R: We're there\
L: I Guess\
R: You wanna try that and see?\
L: Sure\
L: 1B then 3A, then 2,1, then...\
R: 2,1,3.   If we put 2,1,3 and then put 1,2,3,5; move 4,
then put 1,2,3 there and can move 5 there\
L: what now?\
R: We have 2, then 1 and move 3 here, then put 2 and 1 here.
So it's 1,2,3,5. We can put 4 there, then move 1,2,3, and
put 5 there. Think that will work?\
L: Ok\
R: Ok...so\
R: JM 1C.   2A, yeah\
R: 2,1,3\
L: Then 2B?\
R: Ok\
L: 3C?   JM 3C\
L: No\
R: Cause if we go 2B, 1B, we can take 3 off and put 4 there\
L: You do... Oh gosh!   Do 1B, then 3A, no\
R: 1A, What did you say, 1B,3A?\
L: Yeah, then 4 is on the bottom\
R: 2,1,3;   214\
L: Ok, oh ok.   So 1B, then...\
L: 1A, then 2c...no. How are we gonna get 4 on there?\
R: Oh, your right\
L: then 2c, then 1c and...\
R: Or if we move this one here, so we have 4,1,2. I think
we should move 4B\
R: JM 4B.   Then 1B?\
L: Then what?\
R: 1A,2B.   1,2,4,3 and if we could grt 1,2,3,4\
L: ok, yeah\
L: Then do 1c, 2a and...\
R: Ok 1,2,1,3\
L: Yeah\
R: Move 1,2...\

APPENDIX G
EXAMPLES OF CODING SEGMENTS

Coding
Category   Example Verbal Segments

Commands   Move 1 to 2
                        Move 3 on 5
                        Just move 4 over here
                        1 back on 3
                        Ummm, put 1 on C

Explanations   You can probably put, yeah you can move 1 at a
                        time
                        OK now I need to move the 3 on top of the 4
                        We can't put a bigger box on top
                        I was thinking about putting 1 here, cause
                        when we put 1 up here and 2 down

Suggestions   Then maybe 1 on top of 5
                        Why don't you move 3 to A
                        I would put 2 to C so you can move 1
                        Let's move block 1 over to B or C

Acknowledgment   Yeah, OK
                        all right
                        No

Planning   You want to put 1 on 2, then move 3 over and
                        then put 4 here
                        OK, 1 on C, 2 on B, then you wanna say 3 on 4
                        Now if we move 1 to 4, we can move 2 to 3 then
                        1 back to ....

Question   Now you want to move 1 to 2
                        Now what, 2 to B
                        Whose turn is it

Utterance   Ohhhh
                        Ummmm

Other   1 on top of
                        wait if we... Jack can we give up today

68

REFERENCES

Ahlum-Heath, M. E., & Di Vesta, F. J. (1986). The effect of conscious controlled verbalization of a cognitive strategy on transfer in problem solving. Memory & Cognition, 14 (3), 281-285.

Aircraft crash. (1994, Jul 15). Tallahassee Democrat, p. A4.

Barker, J. M., Jr., Clothier, C. C., Woody, J. R., McKinney, E. H., Jr., & Brown, J. L. (in press). Crew resource Management: A simulator study comparing fixed versus formed aircrews. Aviation Space and Environmental Medicine.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1990). Cognitive psychology and team training: Training shared mental models of complex systems. Human Factors Society Bulletin, 33, 1-4.

Cohen, J. & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum Associates.

Crutcher, R. J. (1994). Multiple Protocol Analysis System (Version 1.0) [Computer software]. Chicago: University of Chicago.

Ericsson, K. A. & Simon, H. A. (1993). Protocol Analysis. Cambridge, MA: MIT Press.

Foushee, H. C., Lauber, J. K., Baetge, M. M., & Acomb, D. B. (1986). Crew factors in flight operations: III. The operational significance of exposure to short-haul air transport operations. (NASA Tech. Memo 88322). Mountain View, CA: NASA.

Haynes, A. C. (1991). United 232: Coping with the "one-in-a-billion" loss of all flight controls. Avionics News, pp. 30-61.

Janis, I. L., & Mann, L. (1977). Decision making: A psychological analysis of conflict, choice, and commitment. New York: The Free Press.

Kanki, B. G., & Foushee, H. C. (1989). Communication as group process mediator of aircrew performance. Aviation Space and Environmental Medicine, 60, 402-410.

Leon-Carrion, J., Morales, M., Forastero, P., Dominguez-Morlaes, M., Murillo, F. , Jimenez-Baco, R., and Gordon, P. (1991). The computerized Tower of Hanoi: A new form of administration and suggestions for interpretation. Perceptual and Motor Skills, 73, 63-66.

Morgan, B. B., & Lassiter, D. L. (1992). Team composition and staffing. In R. W. Swezey, & E. Salas, (Eds.), Teams: Their training and performance (pp.75-100). Norwood, NJ: Ablex.

Orasanu, J., & Salas, E. (1992). Team decision making in complex environments. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok, (Eds.), Decision making in action: Models and methods (pp. 327-345). Norwood, NJ: Ablex.

Orasanu, J. (1990). Shared mental models and crew decision making (Tech. Rep. No. 46). Princeton, NJ.

Report blames crew in fighters' deaths. (1994, Aug 21). Tallahassee Democrat, p. A3.

Shaw, M. E. (1976). Group dynamics: The psychology of small group behavior. New York: McGraw-Hill.

Simon, H. A. (1975). The functional equivalence of problem solving skills. Cognitive Psychology, 7, 268-288.

Thordsen, M. L., & Klein, G. A. (1989). Cognitive processes of the team mind. Yellow Springs, OH: Klein Associates Inc.

Woody, J. R., McKinney, E. H., Jr., Barker, J. M., Jr., & Clothier, C. C. (1994). Comparison of fixed versus formed aircrew in military transport. <u>Aviation Space and Environmental Medicine, 65,</u> 153-156.

BIOGRAPHICAL SKETCH

JOHN M. BARKER, JR.

BORN: 5 December, 1958 in Kankakee, Illinois

EDUCATION: M.S. Florida State (Cognitive Psychology), 1990
B.S U.S. Air Force Academy (Human Factors Eng.), 1980

PROFESSIONAL ASSOCIATIONS: Human Factors and Ergonomics Society, Aerospace Medical Association, and Association of Aviation Psychologists

TEACHING EXPERIENCE: 1991-1993 assistant professor U.S. Air Force Academy
1985-1988 instructor pilot U.S. Air Force

RESEARCH EXPERIENCE: Barker, J. M., Jr. (1990). Detecting target words while monitoring multiple auditory inputs. Unpublished master's thesis.

Barker, J. M., Jr., Clothier, C. C., Woody, J. R., McKinney, E. H., Jr., & Brown, J. L. (in press). Crew resource Management: A simulator study comparing fixed versus formed aircrews. Aviation Space and Environmental Medicine.

Derry, S., Weaver, G., Liou, Y., Barker, J. M., Jr., & Salazar, H. (1990). Inducing flexible transfer in novice problem solvers: Effects of three instructional treatments. Paper presented at the American Educational Research Association Conference.

Woody, J. R., Mckinney, E. H., Jr., Barker, J.M., Jr., & Clothier, C.C.(1994). Comparison of fixed versus formed aircrew in military transport. Aviation space and Environmental Medicine, 65, 153-156.

NASA supported Total Quality Management(TQM) study at Cape Canaveral AFS, FL (1992)

NASA team research project at Kennedy Space Center, Fl (1993)

MILITARY EXPERIENCE: Major, U.S. Air Force
Senior pilot (3000 hours) Aircraft flown:
C-141, C-12, UV-18, T-37, & T-38
Aircraft Maintenance officer on F-4D aircraft
Personnel officer Department of Behavioral
Sciences & Leadership U.S. Air Force Academy